AD-A206 990

DTIC ELECTE APR 1 8 1989 S D

DTIC FILE COPY

(2)

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION: unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for public release, distribution unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S): Contract F49620-88-C-0025	5. MONITORING ORGANIZATION REPORT NUMBER(S): AFOSR-TR- 89-0445
6a. NAME OF PERFORMING ORGANIZATION: Yale University; 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION: AFOSR/NE
6c. ADDRESS (City, State, and ZIP Code): 333 Cedar Street, New Haven, CT 06510	7b. ADDRESS (City, State, and ZIP Code): Bldg. 410, BAFB DC 20332
8a. NAME OF FUNDING/SPONSORING ORGANIZATION: AFOSR; 8b. OFFICE SYMBOL (If applicable): NE	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: F49620-88-C-0025
8c. ADDRESS (City, State, and ZIP Code): Bolling AFB, Washington, DC 20332	10. SOURCE OF FUNDING NUMBERS: PROGRAM ELEMENT NO. 61102F; PROJECT NO. 2305; TASK NO. B1; WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification)
A NEURAL NETWORK APPROACH TO MODEL-BASED RECOGNITION

12. PERSONAL AUTHOR(S)
Gene R. Gindi

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Annual	FROM 12/1/87 TO 12/1/88	2/1/89	13

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	

19. ABSTRACT (Continue on reverse if necessary and identify by block number)
The problem of recognizing structured patterns (object recognition) is being pursued on several fronts. The main thrust of the work involves the design of networks that store some notion of a relational model of an object and performs recognition via a version of graph matching. This approach is governed by the use of objective functions to both specify the network and the problem to be solved. The dynamics of the net thus carry out an optimization procedure. Key here is the incorporation into the objective function of compositional and specialization hierarchy of models, and provision to perform dynamic grouping (perceptual organization) of the input data. Results so far show very good performance for versions where data is preprocessed into a form matchable to the database, but poorer performance on more difficult problems where the network must itself organize raw data into relational structures for matching. A related effort explores aspects of "traditional" associative memories that may be of use in more complex networks. Questions of performance, storage and robustness are addressed. A new fast learning algorithm is

(continued on next page)

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: ☒ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION: UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL: Giles	22b. TELEPHONE (Include Area Code): (202) 767-4931 or 32; 22c. OFFICE SYMBOL: NE

DD Form 1473, JUN 86 — Previous editions are obsolete. — SECURITY CLASSIFICATION OF THIS PAGE

89 4 ...

proposed for a CMAC network. Work in optical implementation of some of these networks constitutes a third front. The main problem here is to use optics to form a fixed interconnection network between layers of 2-D nodes (neurons). Two means of using spatial multiplexing to effect a 4-D interconnect between two 2-D node planes are used: multifaceted holograms and multichannel incoherent image systems. Nonlinearities are implemented electrically after detection of the light intensity. Computer simulations and experiments with e-beam fabricated holograms designed to code an associative memory and a winner-take-all network reveal a host of technical problems resulting in a limited achievable accuracy for the connection weight.

AFOSR-TR- 89-0445

One-Year Progress Report
AFOSR Contract F49620-88-C-0025
A Neural Network Approach to Model-Based Recognition

Gene Gindi
Yale University

1 February 1989

1 Summary

In this work, the investigators explored the area of neural-nets for object recognition and also investigated hardware issues in the optical implementation of neural nets. The approach underlying the design of networks was that of optimization theory to both specify the problem and the network for its solution. These networks differ considerably from simple pattern matchers (such as the Hopfield content-addresssable memory) where an iconic version of the pattern itself is stored. Instead, the network implements a form of model matching in which the model base is organized as structured graphs, the input data is organized , again by the optimization procedure, into similarly structured graphs by the neural net, and recognition is accomplished by a form of graph matching. The model base is organized hierarchically in that both compositional (whole-part) and specialization (class membership) notions are captured by sparse matrices which serve as pointer structures in the objective functions. This notion of design allows for a uniform style of addressing both low-level visual problems and more traditional high-level recognition problems. Our main experiments were conducted in visual domains involving the recognition of simple stick figures. One version in which the coding of input data into structured graphs was done as a preprocess achieved considerable success. A more ambitous experiment in which this "perceptual organization" is done as part of the optimization achieved limited results, but served to highlight areas for further progress. Though learning was a subsidiary goal here, progress was made in the areas of CMAC contollers and delineations of learning tasks for the graph matching circuits mentioned above. Also, earlier work involving aspects of more traditional associative memories was continued, with progress achieved in the specification of optimal architectures and objective functions for outer-product style memories. Optical implementations were concerned mainly with the problem of implementing ever more general interconnect patterns. The main problem here is to use optics to form a fixed interconnection network between layers of 2-D nodes (neurons). Two means of using spatial multiplexing to effect a 4-D interconnect between two 2-D node planes are used: multifaceted holograms and multichannel incoherent imaging systems. Nonlinearities are implemented electrically after detection of the light intensity. Computer simulations and experiments with e-beam fabricated holograms designed to code an associative memory and a winner-take-all network reveal so far a host of technical problems resulting in a limited achievable accuracy for the connection weight.

2 Introduction

The following is a summary of progress accomplished over the first year of the two-year grant period. A number of papers, Tech reports, etc. are referenced at the end of this 1-year progress report.

The overall theme of the project was the investigation of issues in the organization of visual memory. The project was pursued at the theoretical level as well as at the level of hardware organization, where issues in the optoelectronic implementation of these memorie were pursued.

This project grew out of earlier work in which an approach to object recognition as an associative memory problem was pursued. The big problem with this earlier approach is that objects must be recognized regardless of changes in position, rotation, scaling, and a host of other deformations. One may store iconic patterns directly in a memory and design such invariances directly into the connection pattern, but the circuits quickly become complicated. We follow the approach taken in traditional computer vision systems and store relational models instead of iconic models into the circuit, and demand that the input data itself be organized into such a relational structure and optimally matched to the "nearest" model. Relational models are designed to automatically capture the desired invariances; the perceptual organization of input data into relational structure proceeds simultaneously with the matching process. Like some associative memory approaches, this one also uses objective functions for specification of the problem and design of the circuit. The optical implmentation aspect is largely independent of the distinction between the associative memory approach and the optimiing graph matching approach. In each case, the main task is to use the capabiliies of optics and electronics to for general interconnect patterns and implemnbt nonlinearities.

Below we list a summary of accomplishments completed during the first year of the grant period. This is follwed by a more detailed discussion.

3 Summary of Completed Research

Below is a succint summary of accomplishments of the first year. Each item is followed by a reference to an appropriate paper or technical report that gives a more detailed account.

- We implemented two networks for simple object recognition and performed analysis and simulation experiments. Each of these successfully matched simple stick figures to a database of models. It was able to find multiple objects and specializations of objects. One version [4] used an unconstrained optimization technique for net dynamics; the other incorporated "Lagrange multiplier neurons" to implement hard constraints [14]. In each case, the difficult task was made tractable by hand coding the input data into relational structures suitable for matching. The resulting objective functions were quadratic and the net worked well.

- A more difficult version of the above task was attempted. Again, the idea is recognize simple stick figures, but now the network, as part of the optimization process, had to group input sticks into potentially meaningful relational structures. This resulted in 5th-order objective functions. While success of the network was limited, much was learned. Results are reported in [8], [9], and [6].

- A version involving recognition of 3D objects was completed. This [15] network recognizes 3D-stick figures from a 2D projection.

- A paper discussing the general approach to object recognition espoused by us and collaborators was invited for publication by the new journal *Neural Computation* [6].

- E-beam multifacetted holograms for optical neural net interconnects were fabricated for an associative memory problem and a winner-take-all problem. The optical results were compared to results from a simulation program written to model sources of error in the e-beam scheme. This resulted in some understanding of the engineering problems associated with the optical interconnect scheme. Results and analyses are shown in this progress report.

- In support of the optical effort, we completed a study on performance of outer-product associative memories. One result [1] showed that versions with self interconnects perform better than those without; another [2] showed that a version with positive only interconnects can be made to perform well. Both of these results have ramifications for optical schemes.

- An additional study on optimal architectures for outer product associative memories was completed [13]. A universal architecture that makes optimally efficient use of hardware is proposed in this study.

- Though learning hasn't been a major theme of the work, some progress was made in initial studies for learning distance metrics in the graph matching networks mentioned above. In an unrelated study, a fast, general purpose, supervised learning algorithm based on CMAC models was devised [11] [12]. On a popular test case, it greatly outperforms backprop.

4 Discussion

The work splits into three categories: networks for model matching, optical implementations, and analyses of associative memory. The optical work was carried out as a subcontract at the University of Arizona and is discussed in a self-contained section.

4.1 Networks for model matching

Neural net tasks for visual recognition is often thought of as a variant of some simple pattern matcher, such as the the Hopfield associative memory or a simple perceptron. These schemes are limited in two ways: "objects" are represented iconically instead of in the more efficient manner of relational structures, and there is no provision for efficiency in search by using notions of hierarchy. Of course, both of these ideas are common in traditional computer vision, but here, we propose a way of incorporating these crucial notions into a neural-net paradigm.

We introduce an optimization approach for solving problems in computer vision that involve multiple levels of abstraction. Specifically, our objective functions can include compositional hierarchies involving object-part relationships and specialization hierarchies involving object-class relationships. The advantage of hierarchical organization is that it makes the search process involved in image interpretation easier to express and more efficient. The large class of vision problems that can be subsumed by this method includes traditional model matching, perceptual grouping, dense field computation (regularization), and even early feature detection which is often formulated as a simple filtering operation. This raises the possibility of solving within a single vision system both low-level and high-level problems in a uniform manner.

Our approach involves casting a variety of vision problems as inexact graph matching problems, formulating graph matching in terms of constrained optimization, and using analog neural networks to perform the constrained optimization. Figs 1 and 2 illustrate the basic idea for a simple graph matching problem.

Our extension of graph-matching to model-based object recognition involves regarding one of the graphs as a "model" graph, which is supposed to represent the knowledge of shapes within the system, and the other graph as a "data" graph which is obtained from the current input data to the system. The model-side nodes are simply called "models" and data-side nodes are called "frames" (denoted F_i), which are collections of analog neurons representing parameters of an object (and are denoted F_{is}, where s indexes the parameters of a single frame). The instantiation of a model in the image is expressed by "turning on" a match neuron $M_{\alpha i}$ between a model α and its matched Frame F_i. We will refer to such a network of frames and models as "Frameville". Basic notions of Frameville are discussed in Figs 3,4,5,6

Exact Graph Match Via a TSP-like Network

Represent a graph by a sparse binary matrix whose ijth element is unity if node i is connected to node j, and is zero otherwise. Given two graphs represented by $G_{\alpha\beta}$ and g_{ij}, introduce a match matrix $M_{\alpha i}$, where $0 < M_{\alpha i} < 1$, to represent the correspondence between nodes α and i .

A simple objective function maximizes the number of local consistency rectangles :

$$E_1(M) = -\sum_{\alpha\beta}\sum_{ij} G_{\alpha\beta} g_{ij} M_{\alpha i} M_{\beta j}. \tag{1}$$

while other terms reflect the constraint of one-to-one matches between nodes:

$$E_2(M) = \sum_{\alpha}(\sum_{i} M_{\alpha i} - 1)^2 + \sum_{i}(\sum_{\alpha} M_{\alpha i} - 1)^2 \tag{2}$$

and limit the values of the match neurons to the desired range:

$$E_3(M) = \sum_{\alpha i} \int^{M_{\alpha i}} g^{-1}(x)dx. \tag{3}$$

where g is a sigmoidal gain function.

Graph matching is performed by a gradient descent procedure on the combined objective $E = E_1 + E_2 + E_3$, or by a constrained optimization technique involving the introduction of Lagrange multiplier neurons.

FIG 1

Exact Graph Matching

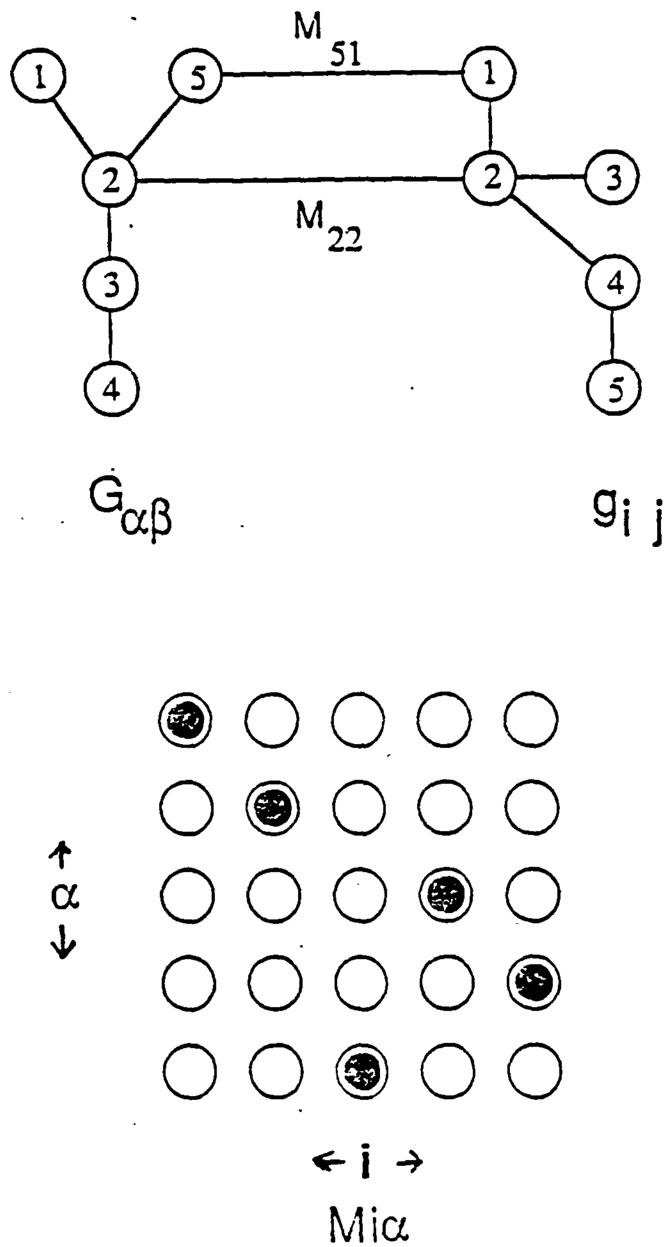


FIG 2
An illustration of a graph matching problem

Introducing Frames

For model matching, $G_{\alpha\beta}$ becomes $INA_{\alpha\beta}$ denoting part-whole relationships among models α and β. Likewise g_{ij} becomes ina_{ij} denoting part-whole relationships between data nodes.

The data nodes themselves become "Frames" which represent visual abstractions in terms of a few parameters. The i'th frame, $\vec{F}_i$, contains "slots" for analog parameters F_{is} and may participate in part-whole relationships (ina_{ij}, now variable) with other frames $\vec{F}_j$. The process of dynamically employing a frame to represent an abstraction required by data is called "allocating" the frame.

The graph-matching objective $E_1(M)$ again maximizes the number of consistency rectangles, but consistency now involves the variable ina links and model-specific criteria H:

$$E_1(M, ina, F) = -\sum_{\alpha\beta}\sum_{ij} INA_{\alpha\beta}\, ina_{ij} M_{\alpha i} M_{\beta j} H^{\alpha\beta}(\vec{F}_i, \vec{F}_j). \qquad (4)$$

FIG 3

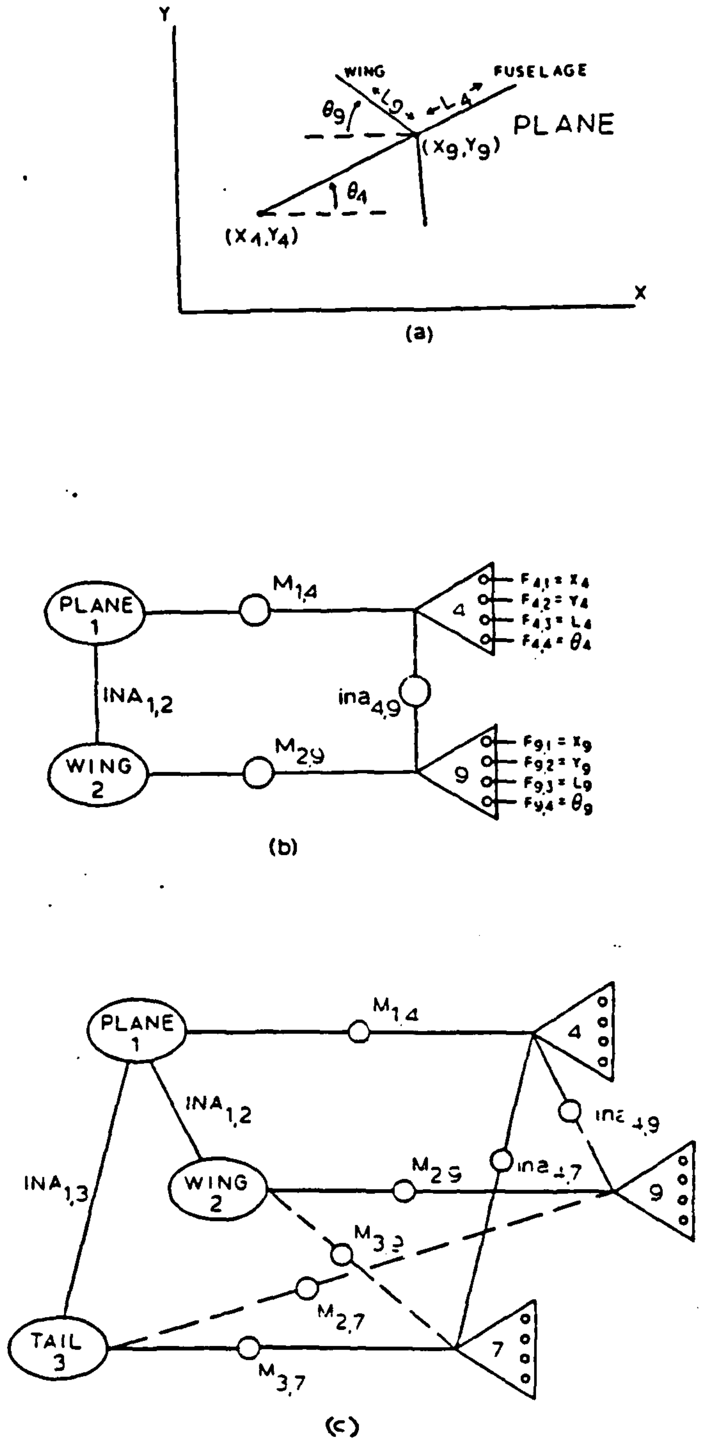


Figure 2: **Example of Frameville rectangle rule.**(a) Example object, a plane, consists of parameterized parts fuselage and wings. (b) Shows the rectangle relationship between frames (triangles) representing a wing and the fuselage of a plane. Circles denote dynamic variables, ovals denote models, and triangles denote frames. For the plane and wing models, the first few parameters of a frame are interpreted as position, length, and orientation. (c) Shows the sibling competition among parts. The match variables along the dotted lines ($M_{3,9}$ and $M_{2,7}$) are suppressed in favor of those along the solid lines ($M_{2,9}$ and $M_{3,7}$).

FIG 4

FRAMEVILLE MATH

Objective:

Data and models must be consistent ...

$$E_1(M, ina, F) = -\sum_{\alpha\beta}\sum_{ij} INA_{\alpha\beta} ina_{ij} M_{\alpha i} M_{\beta j} H^{\alpha\beta}(\vec{F}_i, \vec{F}_j). \quad (5)$$

Constraints:

Unique matches for the parts of an object ...

$$\sum_{\alpha} INA_{\alpha\beta} M_{\alpha i} - \sum_{j} ina_{ij} M_{\beta j} = 0 \quad (6)$$

$$\sum_{i} ina_{ij} M_{\alpha i} - \sum_{\beta} INA_{\alpha\beta} M_{\beta j} = 0. \quad (7)$$

Unique specialization through discrimination tree ...

$$M_{\alpha i} - \sum_{\beta} ISA_{\alpha\beta} M_{\beta i} = 0. \quad (8)$$

M and ina make decisions ...

$$\begin{aligned} M_{\alpha i}(1 - M_{\alpha i}) &= 0 \\ ina_{ij}(1 - ina_{ij}) &= 0. \end{aligned} \quad (9)$$

(or use standard analog gain term.)

FIG 5

Introduction of a Specialization Hierarchy

Indexing into a large database of models may be made efficient by the introduction of a specialization hierarchy. We index the models (and the database of metrics $H^{\alpha\beta}$) by introducing a static graph of pointers $ISA_{\alpha\beta}$ to act as both a discrimination network and an inheritance hierarchy. Note that property inheritance is automatically achieved by allowing the same frame to match to a model and just one of its specializations:

$$M_{\alpha i} - \sum_{\beta} ISA_{\alpha\beta} M_{\beta i} = 0. \tag{10}$$

The additional verification of properties specific to the specialization is simply expressed as additional model-specific constraints involving the parameters.

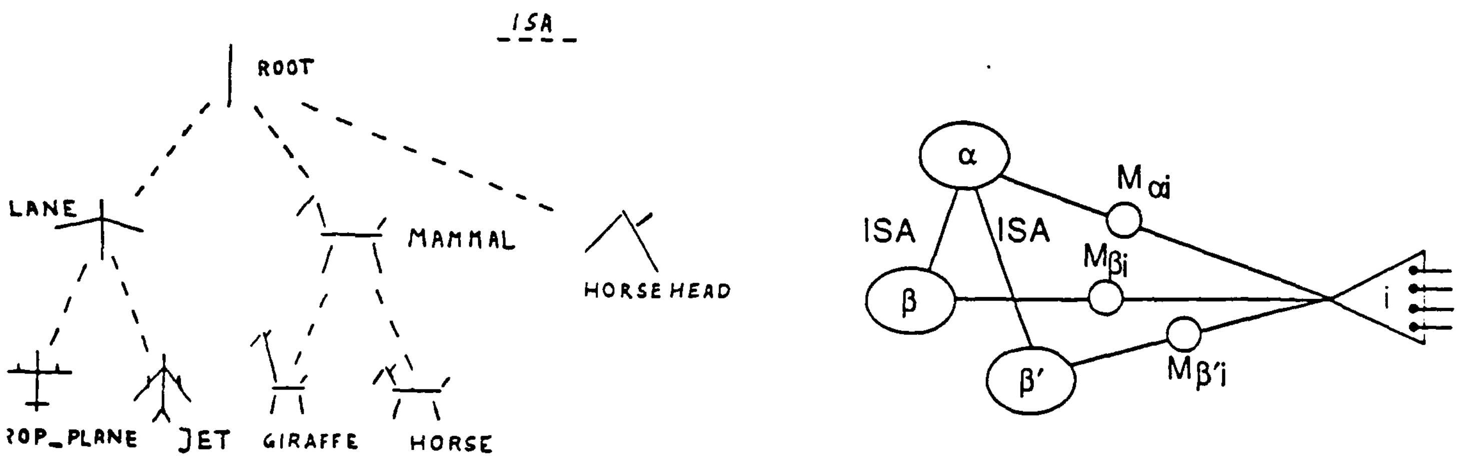


FIG 6

The incorporation of a compositional hierarchy and a specialization hierarchy on the model side is achieved via graph-arcs called *INA* links and *ISA* links respectively. The objective function includes terms representing a simultaneous match of a model to an object (on the data side) and the parts of the model to parts of the object in a consistent fashion. An objective function may be inherited through the *ISA* links from a model to its specializations and there may be an incremental objective function for each of the specializations. Numerical parameters are represented by using analog neurons and the verification of metrical relationships involving these parameters is achieved by corresponding consistency terms in the objective function.

In order to perform perceptual organization, the data-side compositional hierarchies must be dynamic. To achieve this we introduce dynamic *ina* links on the data side. The *ina* links connect more and less abstract frames, and their evolution corresponds to a search for a perceptual organization consistent with the model graph. Specialization is implicitly achieved by the simultaneous match of a frame to a model and to one of its *ISA* specializations.

The next few figures show experimental results. We note that the "Stickville" simulations of figs 7 and 8 do not involve grouping the input sticks into data side graphs; this is done by hand to make the task easier for the network. The resulting objective functions are quadratic. As seen in figure 8, data in the form of a simple stick figure of a jet is presented to the network. Models of mammal, plane, jet are stored. Both data and object relations are encoded in the connection strengths of the network; in particular $H^{\alpha\beta}(F_i, F_j)$ equals a number measuring the quality of match of data items i and j given that i is identified with model α and j with β. Shown is the state of the match matrix at intermediate and final states. Note that both plane, and its *ISA* specialization (jet) have been found.

The next series of experiments is depicted in Figs 9,10,11 and is described in extensive detail in [8]. The parameter evaluations are not precomputed as before, but are computed dynamically. Translation invariance is achieved by using a specified form of the metric. One unsurprising result is that the network must be given a hint in the form of an initial state in which an abtraction frame containing the parameters of a main part is already successfully matched.

Extensive experimentation and analysis continues with this domain. A central theme that is pursued here is the following: How can the database be organized so that the search process is more successful? These and other aspects are duscussed in [8].

The networks discussed here were hand designed in that both the model base and the definition of the match metric $H^{\alpha\beta}$ were chosen in an *ad hoc* manner. It may be possible to improve the performance of a Frameville network by supervised learning. Learning the database is quite difficult, but improving $H^{\alpha\beta}$ for greater discrimination may be possible. A possible strategy would be to present the network with examples of fully matched models (all match variables set). The match metric $H^{\alpha\beta}$ is then

Experimental Results: Stickville

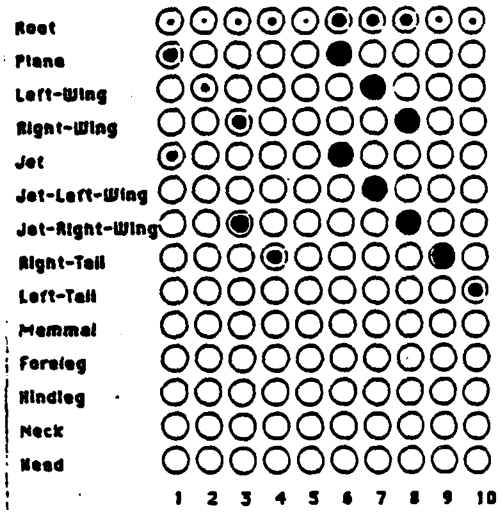
We attempt recognition in Stickville, a simple domain of connected assemblages of linear "sticks". There are severe restrictions here:

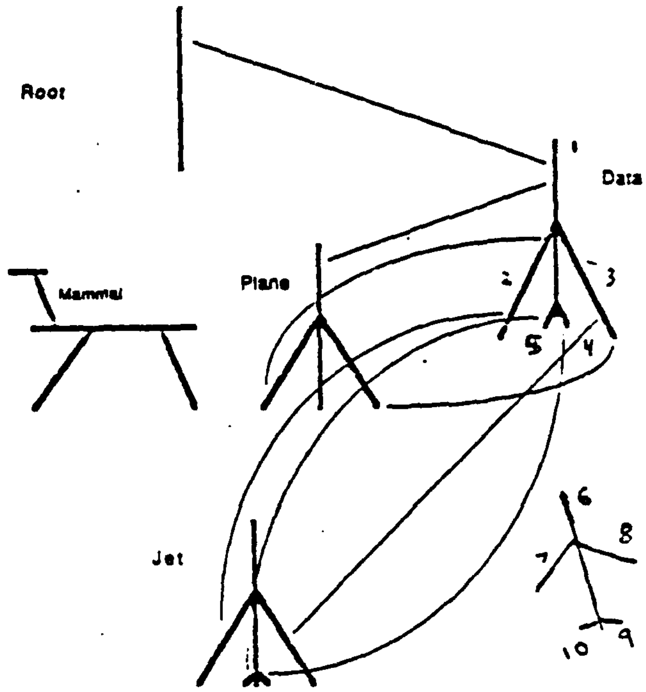
- The data groupings are precomputed, so ina_{ij} are constant.
- The ina matrix represents a tree. The links are undirected.
- Parameters are precomputed. For an attached pair of sticks ($ina_{ij} = 1$), we use relative size, angle, and location of attach point.

Therefore, the only dynamic variables are the match neurons M_{ai} and the objective function is quadratic.

The figure below shows experimental results.

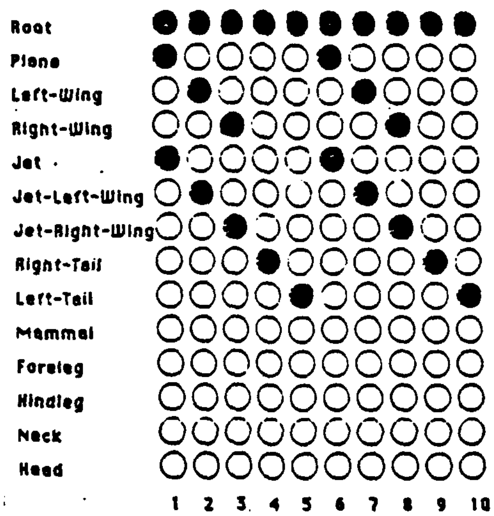
FIG 7





(a)

Stickville match matrix after 28 time steps. The circles represent the dynamic variables whose value is encoded by the radius of the shaded portion. The different columns represent the different data sticks, and the different rows represent different models. The model-base depicted consists of plane, jet, mammal and their parts, and a root model. The data consists of two jets. The unconstrained (Hopfield) update was used.



Stickville match matrix achieves the correct fixed-point after 70 time steps. Note that a stick matches to a model and all of its generalizations (more than one meuron on in a column) and that both the jets have been found (more than one neuron on in a row). The unconstrained (Hopfield) update was used.

FIG 8

Experimental Results: TLville

We experimented with the full Frameville machinery in a character recognition task. The model base is a two-level compositional hierarchy with characters ("T" and "L") composed of unit-length segments in the manner of a standard LED display.

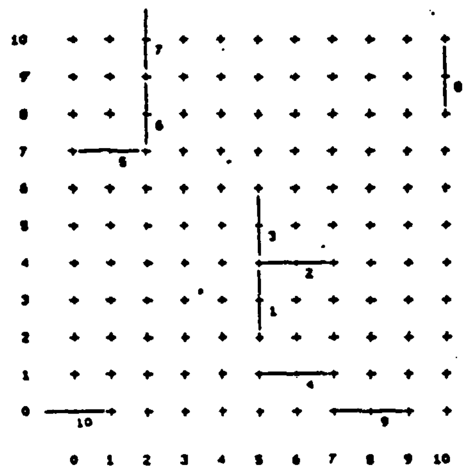
- Frames contain three slots: the x, y, θ coordinates of the segment which the frame represents.
- Abstraction in TLville: High-level frames (those matched to an entire character) contain slots for coordinates of a designated main part.
- Terms in $H^{\alpha\beta}$ like $(x_i - x_j - \Delta x^{\alpha\beta})^2$ enforce translation invariance through analog computation. Other invariances may be similarly implemented.
- No *ISA* specialization mechanism.
- To augment equations 6 and 7, we add penalty terms

$$\sum_{i\alpha} \sum_{\alpha' \neq \alpha} M_{\alpha i} M_{\alpha' i}, \quad \sum_{j\beta} \sum_{\beta' \neq \beta} M_{\beta j} M_{\beta' j}, \quad \sum_{ji} \sum_{i' \neq i} ina_{ij} ina_{i'j}. \tag{11}$$

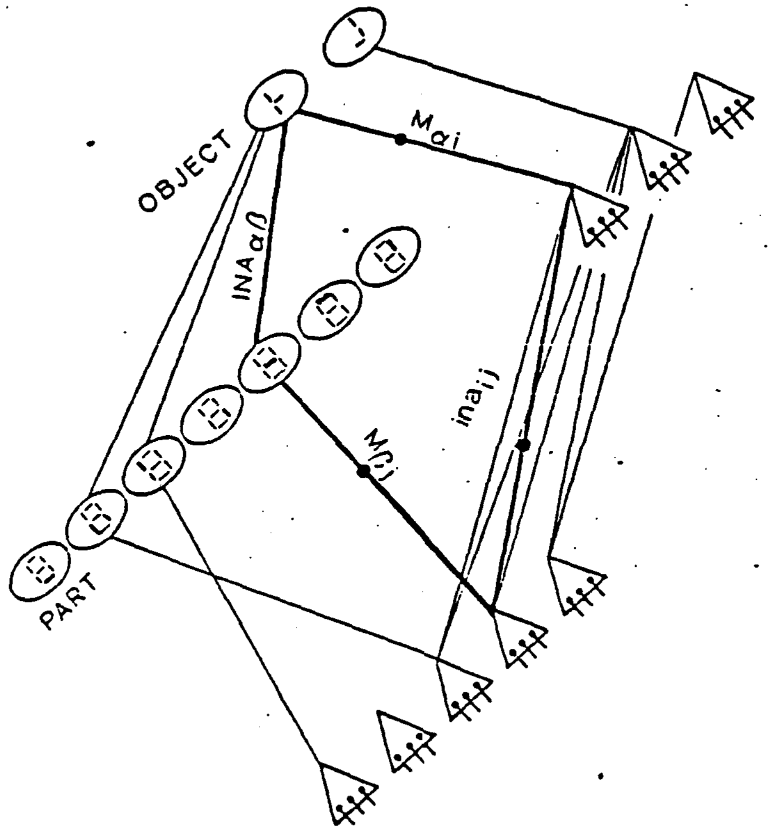
Thus the dynamic variables are M, ina, and F and the resulting objective functions are of order five.

The following figures show experimental results.

FIG 9



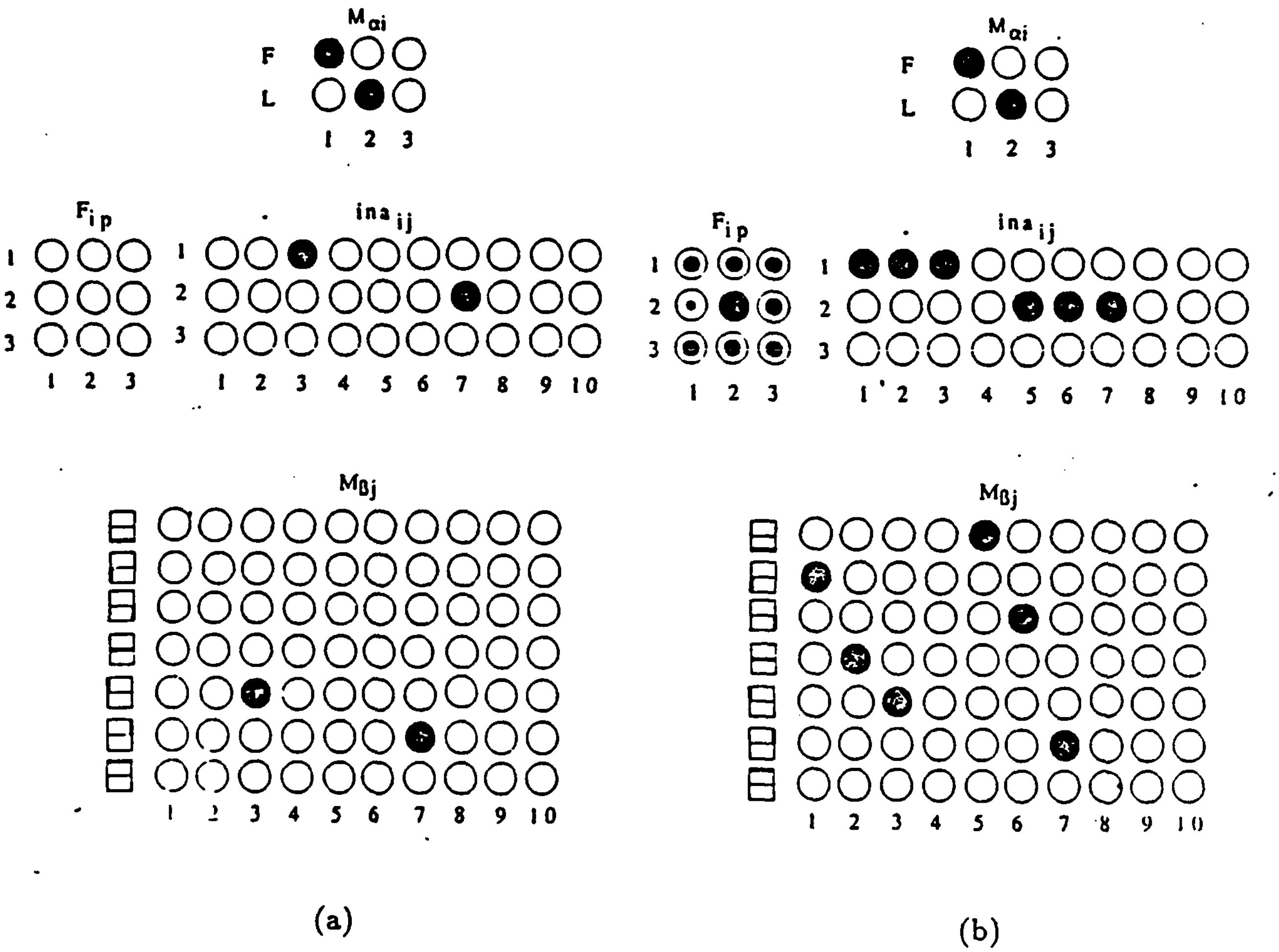
(a)



(b)

(a)**Input data** consists of unit-length segments oriented horizontally or vertically. The task is translation-invariant recognition of three segments forming a "T" junction (e.g. sticks 1,2,3) or an "L" (e.g. sticks 5,6,7) amid extraneous noise sticks. (b)**Structure of network.** Models (represented by ovals) occur at two levels. *INA* links are shown for a "T". Each frame (represented by a triangle) has three parameters: two position coordinates and one orientation coordinate. Also shown are match and *ina* links. The bold lines highlight a possible consistency rectangle.

FIG 10



Match network. The value of each dynamical variable is displayed as the relative area of the shaded portion of a circle. The matrix $M_{\beta j}$ indicates matches between models and frames at the low level; and $M_{\alpha i}$ indicates matches at the high level. The *ina* matrix indicates grouping of ten low-level frames into three high-level frames. The parameters of the high-level frames are displayed in the matrix F_{ip} of linear analog neurons. The parameters of the low-level frames are held fixed and are not displayed. (a) Initial state of network: All neurons set to small random values except those corresponding to matching of main parts. (b) Final state corresponds to successful match of the two junctions and their parts.

FIG 11

changed. For example, it may be desirable to make its value zero or negative for positive training examples and have it assume a large positive value for negative examples. The measure governing the change of H might be the separability of the clusters of energy values reported by the analog computation term for positive and negative examples. The match metric itself could be modified by representing H in parameterized form and descending in the parameter space.

4.2 Work on associative memory and learning

The model matching networks described above differ from the more familiar associatve memories well known to neural netters, but since associative memory serves as a test problem for optical implementaions, several problems in analysis presented themselves. In addition, we continued some previous work in new types of associative memory.

It turns out that a Hopfield style ACAM (Associative Content Addressable Memory) may be implemented optically as projection of a binary input state onto a set of stored memories to obtain a set of inner products, followed by summation of stored memories weighted by the innner products, followed by thresholding. With this kind of optical architecture, it turns out to be convenient to use connections that are positive only instead of bipolar, and to not restrict onesself to eliminating self connections from the network. These two resrictions (bipolar nodes, non self connected nodes) are present in the Hopfield model, but in two papers [1] [2] we show how to remove them. In particular, we show through statistical arguments that the model with self-connected nodes actually works better than the one with without these, and that an all-positive network is possible if the threshold point is selected judiciously.

While the outer-product memories have been well studied and implemented optically, there is reason to prefer alternative models that use simple template matching in conjunction with a layer of internal decision units which compete to perform a winner-take-all (WTA) function. We refer to this a a unary model. There are 3 reasons for our interest: WTA networks are modules of the model-match networks described earlier, they constitute an implementation challenge for the optical effort (see next section), and unary models are interesting in their own right.

With this latter reason as motivation, we completed a study of unary models [13] and showed the following: We present a universal architecture for standartd auto-associative memory models which makes optimally efficient use of hardware. This architecture is described by a bilinear energy functiom. Bot outer-product and unary models can be viewed as special cases of this unversal architecture. The universal architecture uses only the minimal number of binary connections required by information theory to encode the stored memories. For higher order outer product memories, the need for large numbers of internal "product units" is eliminated.

In the previous section, an approach to learning in Frameville was discussed but no committment to a given learning algorithm was mentioned. Reseach on learning, though somewhat peripheral to the immmediate goals in this contract, was conducted. In [11] [12], a learning algorithm in whuich a system learns to approximate mappings by constructing an interpolating lookup table on a lattice of points in the input space.

4.3 Work on optical implementation

This section describes work on optical implmentation performed as a subcontract at the University of Arizona. This progress report was written separately by the subcontract investigators. Figure numbers refer to figures within this report.

AFOSR Progress Report

Objective:
The objective of this portion of the research program is to investigate and develop the implementation of neural-network architectures via optical technology. The strength of optics for this application lies in its ability to perform communication, and so the primary effort is directed at implementation of the massive interconnections required of neural network.

Approach:
The general approach to model-based recognition of visual patterns is that of optimization via a neural network. In the simplest formulation of this type of problem, the interconnection strengths are known and fixed. Thus, it is feasible to implement the interconnections via a fixed nonadaptive structure. In terms of optical imlementation, this fixed structure is easier to realize than an adaptive one, and so the present effort is directed at exploration of the capabilies and limitations of fixed optical interconnects. A fundamental issue concerns the number of nodes and the number of interconnections that will be required for realistic problems. We have chosen to consider nodes layed out on a 2-D array to maximize the number of nodes that can be realized. The associated 4-D interconnection matrix is implemented via spatial multiplexing, essentially producing a 2-D array of 2-D interconnections all ecoded on one large 2-D media.

Two methods of optical interconnection are being investigated, multifacetted holograms and multichannel incoherent imaging systems. In the former, a hologram, hereafter referred to as a subhologram, is dedicated to each node to encode the connection pattern from that node to all the other nodes in the system. A 2-D array of these subholograms forms the composite hologram. The use of Fourier transform holograms allows a simple encoding, where communication to a particular node is associated with a particular spatial frequency (grating) in the subhologram and is independent of the subhologram's spatial position in the composite hologram. The hologram can be designed via any of a number of binary encoding schemes and is fabricated via e-beam lithography.

In the multichannel imaging system approach, there are two geometries. In one, like the subhologram idea, a subtransparency encodes the connection pattern from a particular node to the entire node array, and an array of these, one for each node, meshed together in a 2-D array forms the composite transparency. An array of imaging systems, one for each node, project the individual subtransparencies to the common node input plane. This array of imaging systems can be implemented either via a lenslet array or via a simple shadow-casting optical system. In the other geometry, the multichannel imaging system (lenslet array) forms a spatially replicated image of the node array. Each subtransparency encodes the connections from all the nodes to a particular node. Again an array of these subtransparencies forms the composite transparency. Integration of the light emerging from a subtransparency represents the value of a node input.

Implementation of the nonlinear node function directly via optics is difficult. A dedicated electro-optical node plane, consisting of an array of optical detectors, nonlinear amplifiers, and light sources, appears to be the best approach. The light sources can be either directly driven laser diodes or a spatial light modulator. Difference amplifiers can be incorporated in the electronic chain if bipolar signals are required. This type of structure is compatible with any of the optical interconnection methodologies being considered. In our work, the node function is performed via video detection, digital processing of the digitized video signal, and display on a video monitor. For holographic interconnection, a liquid crystal light valve is used to convert the video display into a coherent amplitude-modulated plane-wave.

Accomplishments:
A simple associative memory problem has been defined to test the operation and performance of the holographic interconnection method. Three states, an A, a B, and an , are stored in a simple autoassociative memory structure. The node field consists of an 8x8 array of binary (0,1) valued nodes. The connection hologram consists of an 8x8 array of subholograms, each encoding the 64 interconnections from its node to the other nodes. Networks with both bipolar and unipolar connection strengths were investigated. A detailed statistical analysis reveals that a system with unipolar interconnections can actually perform better than one with bipolar interconnections (see preprint "Statistical Performance of Outer-Product Associative Memory Models" submitted to Applied Optics).

A winner-take-all (WTA) network was also designed. This network finds the maximum of 16 analog inputs. Again an 8x8 field of nodes is used. Sixteen nodes are used as input nodes, sixteen are used as output nodes, and the rest are used to represent the result of two-input comparisons. Ideally, the maximum analog value should appear on the corresponding output node and all other output nodes should be zero. The interconnections for this type of network are sparse and consists of strictly +1 and -1 connections. On the other hand, the analog node values and the interconnection strengths are required to be very accurate for proper functioning of the network. These two problems, the associative memory problem and the WTA problem, place very different requirements on the network hardware.

A hologram was designed and fabricated to implement the interconnections of the two network problems described above. There are several steps involved in designing the hologram. For bipolar connection patterns the connection matrix is split in two parts, the positive part and the negative part. Each must be encoded in a seperate hologram, the positive part directly and the negative part with a sign reversal so that it too is positive. The unipolar connection pattern will be generated by the hologram as an optical intensity distribution in the output plane of the optical system. The hologram, however, produces an optical amplitude in the output plane. The square root of the connection pattern sets the magnitude that should be produced. The phase, however, is arbitrary and is a degree of freedom inherent to the design process. The hologram transmission should encode the Fourier transform of the complex optical amplitude in the output. This Fourier transform is in general a complex-valued function. The hologram transmission, however, must be binary valued if it is to be fabricated by e-beam lithography. A number of standard techniques exist for this encoding process. All essentially convert the complex function into a real function by spatial carrier modulation, convert the real function into a positive real function by addition of a bias, and encode the sampled magnitude of the positive function by area weighting of the transmissive portion of the hologram. The hologram produces the desired result in a spatially offset region of the output plane with a large on axis bias term. Our particular hologram was encoded as a binary transmission pattern using a technique described by Dallas[1].

One of the major issues for optical interconnects is the achievable accuracy. As a means of investigating this issue, a number of individual test subholograms where also fabricated on the e-beam mask. One of the test holograms encoded the connection pattern dipicted in fig. 1. This connection pattern was used extensively in evaluating the performance of the hologram.

The interconnection pattern produced by a computer-generated hologram is never perfect. All encoding methods result in some degree of error in the resulting connection pattern. Also, the spatial resolution of the e-beam system sets a limit to the dynamic range that can be achieved via the area encoding. This too affects the accuracy of the result. The phase function of the optical amplitude, chosen during the hologram design process, influences the distribution of energy in the hologram and the effective utilization of the limited dynamic range. Holograms with both zero phase and random phase were encoded. Very severe problems can result when the finite extent of the hologram causes a smearing in the output plane. With random phase coding, the smearing can cause interference effects that lead to large error. The smearing effect also causes the energy from the strong central bias to spread into the area where the interconnections are being made. The optical distribution in the output plane produced by the zero phase and the random phase holograms are shown in fig. 2a an 2b respectively. Neither result is particularly good. Clearly, there is a problem with the smearing from the central bias. The random phase hologram produces a higher diffraction efficiency but at the expense of increased error.

Computer simulation of the hologram performance corroborates the story told above. The computer simulation was done by an exact Fourier transform of the computer-generated hologram. A sampled version of the output is shown in fig. 3, where fig 3a corresponds to the zero phase hologram and fig. 3b to the random phase hologram. These can be compared directly to the results in fig. 2; good agreement is noted. This is encouraging for it allows the computer simulation approach to be used to test alternative hologram designs without resort to expensive and time-consuming fabrication. For example, one solution to the problem of smeared light from the on-axis bias light is to make a binary phase hologram from the binary transmission hologram. A binary phase hologram has two transmission values of 1 and π rather than 1 and 0. If there are equal areas of the two transmissions, the on-axis bias signal is identically zero. A simulation of this hologram produces the result shown in fig. 4; a clear improvement over the binary tranmission case

1. B. R. Frieden, ed., The Computer in Optical Research, Springer-Verlag, 1980.

We have attempted to make binary phase holograms by bleaching photographic copies made by contact printing. This approach has not met with much success. It is very difficult, with bleached holograms, to obtain uniform phase independent of spatial frequency.

The full associative memory network was attempted using the electro-optic node plane and the holographic interconnect. The errors in the connection pattern resulted in poor performance for the associative memory problem. The encoded states of the network were not fully stable, though the patterns were stable enough to produce recognizable characters. We are still investigating the behavior of these optically interconnected networks.

Future Directions:
Several areas will be investigated in an attempt to improve the performance of the holographic interconnects. Apodization of the subholograms will be explored as a means of confining the energy and reducing the deleterious effects of smearing. Also, increasing the seperation of the nodes and the seperation of the interconnect region from the central bias can be used to improve performance. There is an obvious tradeoff here between the accuracy of the interconnection and the number of interconnects that can be realized. The issue of accuracy in network interconnections is being studied theoretically in an attempt to delineate exactly what the accuracy requirements are for various kinds of neural networks.

Further investigation of the what phase function should be specified for the ouput amplitude will be done. It is certainly possible that some deterministic phase function is preferable to either the zero phase of the random phase solutions. We will also explore the performance of various hologram encoding procedures including the Lohmann technique, the Lee technique, and the Burch technique. A related idea is to pose the encoding as an optimization problem and employ a technique such as simulated annealing in an attempt to find the "best" solution. This approach might be considered as a network solution to the problem of defining a holographic interconnect for another network problem.

All of the above investigations will be done through computer simulation When an optimal approach is defined, a reasonably large network consisting of perhaps 64x64 nodes will be attempted.

Another goal will be to investigate methods of producing more accurate binary phase holograms. A number of techniques for this exist. We will investigate a photolithographic technique, a photolithographic technique with vacuum deposition of a dialectric coating, and, if possible, photolithography with reactive ion etching.

Investigation of the incoherent optical interconnection methods will also be done in the coming year. A high-resolution film writer using laser scanning technology will be available shortly. Production of transparency masks via this instrument should be relatively straightforward.

Figure Captions:

Fig. 1. Pattern of interconnection of the test subhologram. The connection to the node in the lower left has a value of 5; all other connections are either 0 or 1.

Fig. 2. Holographic interconnection pattern of the test holograms - zero phase (a) and random phase (b). Note the severe spreading of the zero bias light into the interconnection region.

Fig. 3. Calculated interconnection pattern of the test holograms - zero phase (a) and random phase (b). Good agreement with the optically generated interconnection patterns (fig. 2) is noted, apart from the overall photographic intensity.

Fig. 4. Calculated interconnection pattern of a binary phase hologram. Comparison to the results of the binary amplitude hologram (fig. 3a) reveals greatly reduced error from the bias light and improved diffraction efficiency.

FIG 2a

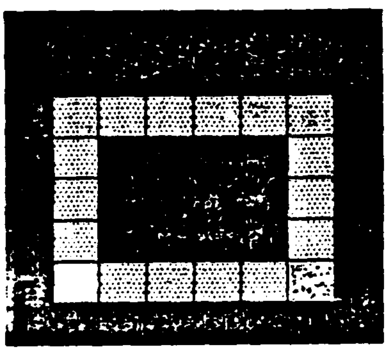


FIG 1

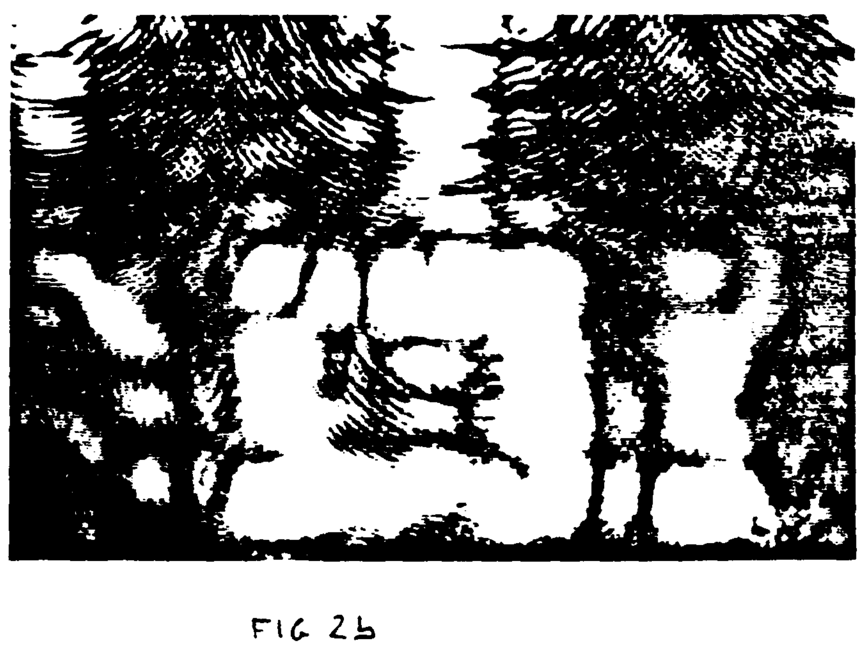
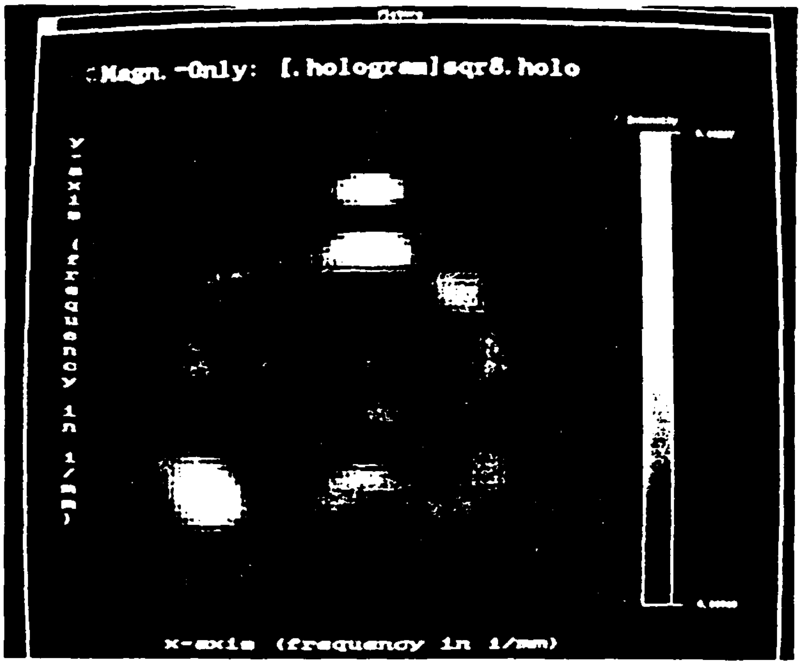
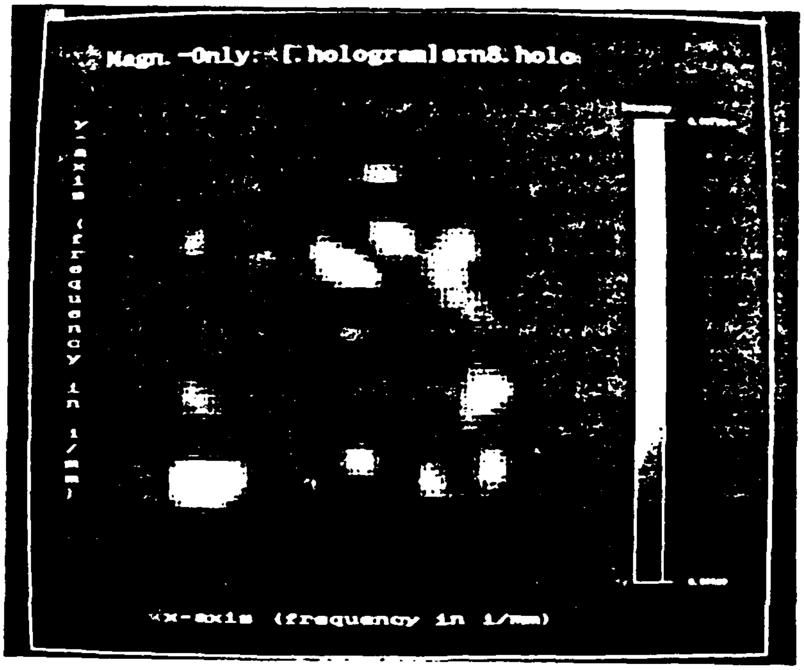


FIG 2b


Magn. -Only: [.hologram]sqr8.holo
y-axis (frequency in 1/mm)
x-axis (frequency in 1/mm)

(a)


Magn. -Only: [.hologram]srn8.holo
y-axis (frequency in 1/mm)
x-axis (frequency in 1/mm)

(b)

FIG 3

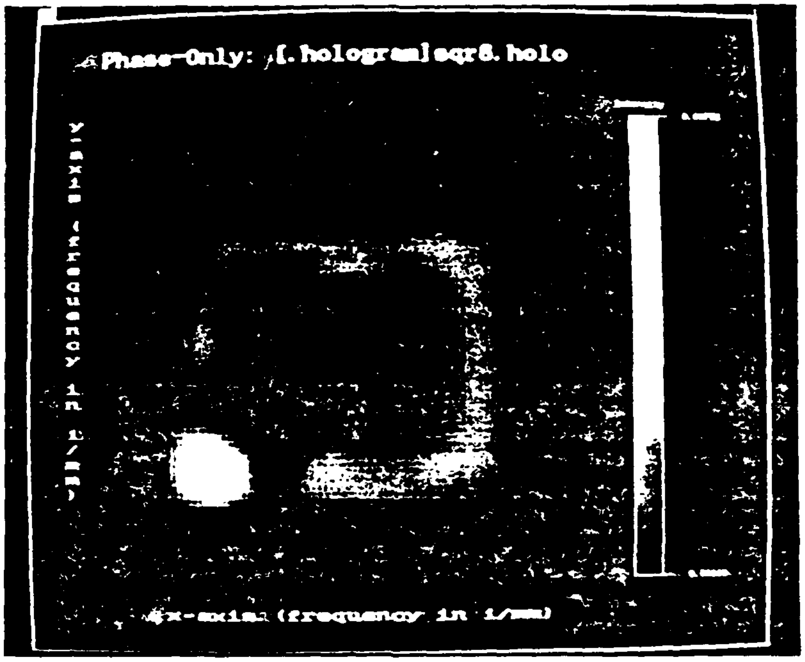

Phase-Only: [.hologram]sqr8.holo
y-axis (frequency in 1/mm)
x-axis (frequency in 1/mm)

FIG 4

5 Summary of Personnel and Publications

Below is a list of professional personnel associated with this grant:

- G.R. Gindi, Associate Prof. Diagnostic Radiology and Electrical Engineering,Yale U.
- A.F. Gmitro, Assistant Prof. Diagnostic Radiology and Optical Sciences, U. of Arizona
- John Moody, Associate Research Scientist, Depertment of Computer Science, Yale University
- Joachim Utans, Graduate Student, Department Electrical Engineering, Yale U.
- Kannan Parthasarathy, Graduate Student, Department of Electrical Engineering, Yale U.
- Paul Keller, Graduate student , Optical Sciences Center, University of Arizona.
- Tony Zador, Graduate Student, Section of Neuroanatomy, Yale U.
- Grant Shumaker, Medical Student, Yale University School of Medicine

Below is a list of publications resulting from this grant:

- G.R. Gindi, A.F. Gmitro, and K. Parasarathy, "Hopfield Model Associative Memory with Nonzero Diagonal Terms in Memory Matrix", Applied Optics, **37**, 129-135 (1988).
- A.F. Gmitro, P. E. Keller, and G.R. Gindi, "Statistical Performance of Outer-Product Associative Memory Models", Applied Optics, (accepted for publication)
- E. Mjolsness, G. R. Gindi, P. Anandan, "Neural Networks for Model Matching and Perceptual Organization", 1988 Neural Information Processing Systems. (to appear in book of proceedings)
- P. Anandan, E. Mjolsness, G.R. Gindi, "Low-level Visual Grouping via Optimization in Neural Networks", submitted to 1989 Conference on Computer Vision and Pattern Recognition.
- T. Zador, G. Gindi, E. Mjolsness, P. Anandan, "Neural Network for Model based Recognition - Simulation Results", Proc. Intl. Conference on Neural Networks, p.532, Boston 1988.

- E. Mjolsness, G. R. Gindi, P. Anandan, "Objective Functions for Visual Recognition: A Neural Network that Incorporates Inheritance and Abstraction", Proc. Snowbird Conference on Neural Networks for Computing, Salt Lake City, Utah 1988.

- E. Mjolsness, G.R. Gindi, and P Anandan, "Optimization in Model Matching and Perceptual Organization" to appear in *Neural Computation* Vol. 1 no. 2 May 1989

- E. Mjolsness, G.R. Gindi, and P Anandan, Objective Functions in Recognition",AAI Spring Symposium on on Physical and Biological Approaches to Computational Vision" (Stanford March 1988) pp. 105-107

Below is a list of abtracts and technical reports associated with this grant:

- J. Utans, G. Gindi, E. Mjolsness, and P. Anandan' "Neural Networks for Object Recognition within Compositional Hierarchies", Technical Report, Yale University Center for Systems Science, TR 8903, February 1989.

- E. Mjolsness, G.R. Gindi, and P Anandan, "Optimization in Model Matching and Perceptual Organization" Technical Report YALEU/DCS/RR-634 Yale U. Dept Computer Science, June 1988

- A.F. Gmitro and P. Keller, "Space-Variant Optical Interconnects via Multifaceted Planar Holograms", presented at Optical Society of America Annual Meeting, Santa Clara CA, October 1988.

- J. Moody and C. Darken, "Learning with Localized Receptive Fields", Tech Report YALEU/DCS/RR-649, Yale U.Dept Computer Science, Sptember 1988.

- J. Moody and C. Darken, "Fast Learning in Networks of Locally-Tuned Processing Units", Tech Report YALEU/DCS/RR-654, Yale U.Dept Computer Science, Sptember 1988.

- J. Moody, "Optimal Architectures and Objective Functions for Associative Memory: A Summary" internal report (available on request) Yale U.Dept Computer Science, January 1989.

- G. Shumaker, G. Gindi,"Stickville", Technical Report, Yale University Center for Systems Science, (in preparation)

- G. Shumaker, "A Stickville calable of Rotational Projection: Rotatorberg". Technical Report, Yale University Center for Systems Science, (in preparation)

References

[1] G.R. Gindi, A.F. Gmitro, and K. Parasarathy, "Hopfield Model Associative Memory with Nonzero Diagonal Terms in Memory Matrix", Applied Optics, **37**, 129-135 (1988).

[2] A.F. Gmitro, P. E. Keller, and G.R. Gindi, "Statistical Performance of Outer-Product Associative Memory Models", Applied Optics, (accepted for publication)

[3] E. Mjolsness, G. R. Gindi, P. Anandan, "Neural Networks for Model Matching and Perceptual Organization", 1988 Neural Information Processing Systems. (to appear in book of proceedings)

[4] T. Zador, G. Gindi, E. Mjolsness, P. Anandan, "Neural Network for Model based Recognition - Simulation Results", Proc. Intl. Conference on Neural Networks, p.532, Boston 1988.

[5] E. Mjolsness, G. R. Gindi, P. Anandan, "Objective Functions for Visual Recognition: A Neural Network that Incorporates Inheritance and Abstraction", Proc. Snowbird Conference on Neural Networks for Computing, Salt Lake City, Utah 1988.

[6] E. Mjolsness, G.R. Gindi, and P Anandan, "Optimization in Model Matching and Perceptual Organization" to appear in *Neural Computation* Vol. 1 no. 2 May 1989

[7] E. Mjolsness, G.R. Gindi, and P Anandan, "Objective Functions in Recognition",AAAI Spring Symposium on on Physical and Biological Approaches to Computational Vision" (Stanford March 1988) pp. 105-107

[8] J. Utans, G. Gindi, E. Mjolsness, and P. Anandan "Neural Networks for Object Recognition within Compositional Hierarchies", Technical Report, Yale University Center for Systems Science, TR 8903, February 1989.

[9] E. Mjolsness, G.R. Gindi, and P Anandan, "Optimization in Model Matching and Perceptual Organization" Technical Report YALEU/DCS/RR-634 Yale U. Dept Computer Science, June 1988

[10] A.F. Gmitro and P. Keller, "Space-Variant Optical Interconnects via Multifaceted Planar Holograms", presented at Optical Society of America Annual Meeting, Santa Clara CA, October 1988.

[11] J. Moody and C. Darken, "Learning with Localized Receptive Fields", Tech Report YALEU/DCS/RR-649, Yale U.Dept Computer Science, September 1988.

[12] J. Moody and C. Darken, "Fast Learning in Networks of Locally-Tuned Processing Units", Tech Report YALEU/DCS/RR-654, Yale U.Dept Computer Science, September 1988.

[13] J. Moody, "Optimal Architectures and Objective Functions for AssociativeMemory: A Summary" internal report (available on request) Yale U.Dept Computer Science, January 1989.

[14] G. Shumaker, G. Gindi,"Stickville", Technical Report, Yale University Center for Systems Science, (in preparation)

[15] G. Shumaker, "A Stickville Capable of Rotational Projection: Rotatorberg ", , Technical Report, Yale University Center for Systems Science, (in preparation)